

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 3, 2005, 08:43:50 ; Search time 168 Seconds

(without alignments)
657.501 Million cell updates/sec

Title: US-10-813-549-1
Perfect score: 1380
Sequence: MAISIKTPEDIEKRVAGRL.....CEILTIRKDDTIPAIISHDB 264

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqb, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:
 1: /cgmn_6/prodata/1/pubpaal/US07_PUBCOMB.pep:*

2: /cgmn_6/prodata/1/pubpaal/US06_NEW_PUBCOMB.pep:*

3: /cgmn_6/prodata/1/pubpaal/US06_NEW_PUBCOMB.pep:*

4: /cgmn_6/prodata/1/pubpaal/US07_NEW_PUBCOMB.pep:*

5: /cgmn_6/prodata/1/pubpaal/US10_NEW_PUBCOMB.pep:*

6: /cgmn_6/prodata/1/pubpaal/PCTUS_PUBCOMB.pep:*

7: /cgmn_6/prodata/1/pubpaal/US08_NEW_PUBCOMB.pep:*

8: /cgmn_6/prodata/1/pubpaal/US08_PUBCOMB.pep:*

9: /cgmn_6/prodata/1/pubpaal/US09_PUBCOMB.pep:*

10: /cgmn_6/prodata/1/pubpaal/US09C_PUBCOMB.pep:*

11: /cgmn_6/prodata/1/pubpaal/US10A_PUBCOMB.pep:*

12: /cgmn_6/prodata/1/pubpaal/US10A_PUBCOMB.pep:*

13: /cgmn_6/prodata/1/pubpaal/US10B_PUBCOMB.pep:*

14: /cgmn_6/prodata/1/pubpaal/US10B_PUBCOMB.pep:*

15: /cgmn_6/prodata/1/pubpaal/US10C_PUBCOMB.pep:*

16: /cgmn_6/prodata/1/pubpaal/US10D_PUBCOMB.pep:*

17: /cgmn_6/prodata/1/pubpaal/US10E_PUBCOMB.pep:*

18: /cgmn_6/prodata/1/pubpaal/US10F_PUBCOMB.pep:*

19: /cgmn_6/prodata/1/pubpaal/US11A_PUBCOMB.pep:*

20: /cgmn_6/prodata/1/pubpaal/US11_NEW_PUBCOMB.pep:*

21: /cgmn_6/prodata/1/pubpaal/US60_NEW_PUBCOMB.pep:*

22: /cgmn_6/prodata/1/pubpaal/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1
 ; Sequence 12, Application US/10299867

; Publication No. US20030203408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sympon, Carolyn J.

; APPLICANT: Aurora, Rafeev
 ; APPLICANT: Dotson, Stanton B.
 ; APPLICANT: Frazier, Ronald B.
 ; APPLICANT: Woods, Cynthia L.
 ; APPLICANT: Zakeri, Hamideh
 ; APPLICANT: Zhou, Xianzhi

; TITLE OF INVENTION: Human methionine aminopeptidase type 3
 ; TITLE OF INVENTION: (MetAP-3)
 ; FILE REFERENCE: S3181-01-US

; CURRENT APPLICATION NUMBER: US/10299867
 ; CURRENT FILING DATE: 2002-11-19
 ; PRIORITY NUMBER: US 60/125,119

; PRIORITY FILING DATE: 1999-03-11
 ; PRIORITY NUMBER: US 09/523,263
 ; PRIORITY FILING DATE: 2000-03-10

; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 264
 ; TYPE: PRT
 ; ORGANISM: E. coliI MetAP

; US-10-299-867-12
 Sequence 12, Appli
 Sequence 8, Appli
 Sequence 31, Appli
 Sequence 1, Appli
 Sequence 9, Appli
 Sequence 13, Appli
 Sequence 121, Appli
 Sequence 59, Appli
 Sequence 61, Appli
 Sequence 123, Appli
 Sequence 10, Appli

Query Match 100.0%; Score 1380; DB 15; Length 264;
 Best Local Similarity 100.0%; Pred. No. 2.2e-133;
 Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAISIKTPEDIEKRVAGRLAAEVLEMIEPYVKGVSTGELDRICNDYIVNBOHAVSACL 60

OM protein - protein search, using SW model

Run on:

November 3, 2005, 08:31:04 ; Search time 175 Seconds
(without alignments)

772,508 Million cell updates/sec

Title: US-10-813-569-1

Perfect score: 1380

Sequence: 1 MAISRKTPEDIEKMRVAGRL.....CEILTURKDDTIPATISHDE 264

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03;* 1: uniprot_sprot;* 2: uniprot_trembl;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1380	100.0	264	1 AMPM_ECOLI P07906 escherichia
2	1253	90.8	263	1 AMPM_SALTY P10892 salmonella
3	1198	86.8	263	2 06577B Q66718 yersinia ps
4	1195	86.6	263	2 Q82H67 08z167 yersinia pe
5	1158	83.9	264	2 Q6D8E4 06d8e4
6	1137.5	82.4	265	2 Q7NBPB Q7nbpB
7	946	82.4	267	2 Q9CNH2 Q9cnh2
8	937	67.9	264	1 AMPM_BUCALI P57224
9	934	67.7	268	2 Q7VP53 Q7vp53
10	919.5	66.6	260	2 Q886P4 Q886p4
11	919.5	66.3	264	2 Q7VRB7 Q7vrb7
12	914.5	66.3	260	2 Q8BMIL Q8bmil
13	910.5	66.0	268	1 AMPM_HAEIN Q9HXYL Q9hxyl
14	909.5	65.9	261	2 Q9HXYL Q9hxyl
15	902	65.4	251	1 AMPM_BUCAP Q85K71 Q85k71
16	886.5	64.2	270	2 Q65S27 Q65s27
17	881	63.8	258	2 Q83BV1 Q83bv1
18	880.5	63.8	275	2 Q6LN23 Q6ln23
19	878	63.6	258	2 Q8PKW9 Q8pkw9
20	874	63.3	258	2 Q8PAUS Q8pus
21	853.5	61.8	259	2 Q7QKCO Q7qkco
22	853	61.8	258	2 Q7QKCO Q7qkco
23	852	61.7	259	2 Q87f52 Q87f52
24	852	61.7	263	2 Q8D2G1 Q8d2g1
25	849.5	61.6	266	1 AMPM_BUCCB P083CH7 Q9ap3
26	843.5	61.1	265	Q83CH7 Q83ch7
27	823.5	59.7	280	2 Q8PKV1 Q8pkv1
28	804	58.3	272	2 Q82XJ5 Q82xj5
29	798	57.8	271	2 Q62JC3 Q62jc3
30	798	57.8	271	2 Q63T11 Q63t11
31	796.5	57.7	261	2 Q66e99 Q66e99

ALIGNMENTS

32	796.5	57.7	261	2 Q8ZBR7 Q8zbr7	yersinia pe
33	795	57.6	265	2 Q7NVE5 Q7nve5	chromobacte
34	791	57.3	275	2 Q8XZJ2 Q8xzj2	ralstonia s
35	779.5	56.5	254	2 Q6F552 Q6f552	acetinotbact
36	775.5	56.2	292	2 Q87MD7 Q87md7	vibrio para
37	775	56.2	263	2 Q883J5 Q883j5	pseudomonas
38	765.5	55.5	292	2 Q8DBG2 Q8dbg2	vibrio vuln
39	763.5	55.3	292	2 Q7M1F9 Q7m1f9	vibrio vuln
40	761	55.1	265	2 Q6FAV1 Q6fav1	Q9favl acinetobact
41	758.5	55.0	259	2 Q9JX9 Q9jx9	Q9jx9 neisseria m
42	758.5	55.0	259	2 Q7NVR8 Q7nvr8	Q7nvr8 chromobacte
43	745.5	54.1	263	2 Q73IJ4 Q73ij4	Q73ij4 wolbachia p
44	745.5	54.1	260	2 Q73IJ4 Q73ij4	Q73ij4 bordetella
45	743	53.8	273	2 Q7VYD1 Q7vyd1	Q7vyd1

OM protein - protein search, using sw model

Run on:

November 3, 2005, 08:34:05 ; Search time 40 Seconds

(without alignment)

635.031 Million cell updates/sec

Title: US-10-813-549-1

perfect score: 1380

Sequence: 1 MAISIKTPEDIKMRVAGRL.....CEILTRKDDTIPATIHSDE 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79;*

1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

* SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1380	100.0	264	1 DPECM	methionyl aminopeptidase (EC 3.4.11.18) [validated] - Escherichia coli (strain K-12)
2	1380	100.0	264	2 B88501	Nt[Alternative names: metallooligopeptidase; methionine aminopeptidase; peptidase M]
3	1380	100.0	264	3 B90650	C[Species: Escherichia coli]
4	1258	91.2	264	2 AI0528	C[Date: 30-Sep-1988 #sequence revision 30-Sep-1988 #text_change 09-Jul-2004]
5	1258	91.2	264	2 S12027	C[Accession: A27761; S45233; R64740]
6	1195	86.5	263	2 AB0128	R; Ben-Bassat, A.; Bauer, K.; Chang, S.Y.; Myambo, K.; Boosman, A.J.; Chang, S.
7	937	67.9	264	2 AB4957	J; Bacteriol. 169, 751-757, 1987
8	910.5	66.0	268	2 C64138	A; Title: Processing of the initiation methionine from proteins: properties of the Esche
9	909.5	65.9	261	2 D83189	A; Reference number: A27761; MUID:87109068; PMID:3027045
10	852	61.7	259	1 E82845	A; Molecule type: DNA
11	823.5	59.7	280	2 D82097	A; Residues: 1-264 <FUJ>
12	795.5	57.7	261	2 AE0405	A; Cross-references: EMBL:D26562; NID:9473770; PIDN:BAA05612.1; PID:9473823
13	758.5	55.0	259	2 E82029	R; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C <i>et al.</i> ; Rose, D.J.; Mau, B.; Shao, Y.
14	758.5	55.0	259	2 C8108	Science 277, 1453-1462, 1997
15	739	53.6	260	1 E83001	A; Title: The complete genome sequence of Escherichia coli K-12.
16	723.5	52.4	256	2 AG2324	A; Reference number: A64720; MUID:97426617; PMID:9278503
17	697	50.5	276	2 H81580	A; Molecule type: DNA
18	685.5	49.7	259	2 A71644	A; Residues: 1-264 <BLAT>
19	684.5	49.6	259	2 D97859	A; Cross-references: GB:AE00126; GB:U00096; NID:91786358; PIDN: AAC73279.1; PID:91786364;
20	669.5	48.5	253	2 S75731	R; Roderick, S.L.; Matthews, B.W.
21	660	47.8	278	2 AB2774	R; Submitted to the Brookhaven Protein Data Bank, December 1992
22	660	47.8	278	2 H97553	A; Reference number: A51847; PDB:1MAT
23	659	47.8	276	2 A12341	A; Contents: annotation; X-ray crystallography, 2.4 angstroms, residues 2-264
24	646	46.8	274	2 S76999	R; Roderick, S.L.; Matthews, B.W.
25	634	45.9	285	2 BB7106	Biochemistry / 32, 3907-3912, 1993
26	630	45.7	285	2 T55757	A; Title: Structure of the cobalt-dependent methionine aminopeptidase from Escherichia co
27	617	44.7	285	2 G70885	A; Reference number: A4970; MUID:9322948; PMID:8471602
28	594	43.0	298	2 BB8888	A; Description: annotation; X-ray crystallography, 2.4 angstroms
29	42.9	259	2 CT5410	C[Function: monomer]	

Result No.	Score	Query	Match Length	DB ID	Description
30	583	42.2	248	2 JS0493	methionyl aminope
31	568	41.2	275	2 AH1933	methionine aminope
32	565	40.9	305	2 S75743	methionyl aminope
33	559.5	40.5	387	2 S59390	methionyl aminope
34	557.5	40.4	250	2 F72247	methionyl aminope
35	556.5	40.3	369	2 C86257	probable methionin
36	553	40.1	305	2 E85437	methionyl aminope
37	549	39.8	295	2 G81724	methionine aminope
38	543	39.3	291	2 D7162	probable methionin
39	540	39.1	248	2 T4405	methionyl aminope
40	539	39.1	266	2 A81531	methionine aminope
41	539	39.1	291	2 F86516	methionine aminope
42	539	39.1	291	2 E72008	methionine aminope
43	532	38.6	379	2 T40384	probable methionin
44	488.5	35.4	258	2 F70307	methionyl aminope
45	482	34.9	252	2 AD1660	methionine aminope

* ALIGNMENTS

RESULT 1

DPCM

methionyl aminopeptidase (EC 3.4.11.18) [validated] - Escherichia coli (strain K-12)

Nt[Alternative names: metallooligopeptidase; methionine aminopeptidase; peptidase M]

C[Species: Escherichia coli]

C[Date: 30-Sep-1988 #sequence revision 30-Sep-1988 #text_change 09-Jul-2004]

C[Accession: A27761; S45233; R64740]

R; Ben-Bassat, A.; Bauer, K.; Chang, S.Y.; Myambo, K.; Boosman, A.J.; Chang, S.

J; Bacteriol. 169, 751-757, 1987

A; Title: Processing of the initiation methionine from proteins: properties of the Esche

A; Reference number: A27761; MUID:87109068; PMID:3027045

A; Molecule type: DNA

A; Residues: 1-264 <FUJ>

A; Cross-references: EMBL:D26562; NID:9473770; PIDN:BAA05612.1; PID:9473823

R; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C*et al.*; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A; Title: The complete genome sequence of Escherichia coli K-12.

A; Reference number: A64720; MUID:97426617; PMID:9278503

A; Status: nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-264 <BLAT>

A; Cross-references: GB:AE00126; GB:U00096; NID:91786358; PIDN: AAC73279.1; PID:91786364;

R; Roderick, S.L.; Matthews, B.W.

Biochemistry / 32, 3907-3912, 1993

A; Title: Structure of the cobalt-dependent methionine aminopeptidase from Escherichia co

A; Reference number: A4970; MUID:9322948; PMID:8471602

A; Description: annotation; X-ray crystallography, 2.4 angstroms

C[Function: monomer]

A; Description: catalyzes hydrolysis of amino-terminal methionine from proteins

C[Superfamily: aminopeptidase; cobalt; metalloprotein; monomer; protein biosynthesis

Copyright GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 18:02:50 ; Search time 2355 Seconds
(without alignments)

TITLE: TTS-10-013-E10-1
4267.677 Million cell updates/sec

Perfect score: 1380
Sequence: 1 MAISIKTPEDIKMRVAGRL.....CEILTLRKDDTIPALIISNDE 2644

Scoring table: BLOSUM62

Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Command line command: Listing first 45 summaries

-MODEL=frame+_p2n.model -DEV=xlh
-Q=cgn2_1/USPTO_spool/US10813549/runat 03112005 075935 13175/app query.

```
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10813549 @CGN 1_1_3437 @runat_03112005_075935 13175 -NCPU=6 -ICPPU=
```

```
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPPOP=10 -XGAPEXT=0.5 -FGAPEXT=7 -YGAPPOP=10 -YGAPEXT=0.5 -DELOP=6 -DETEXT=7
```

Database : EST:*

```
2: gb->est2:*
3: gb_htc:*
```

4.. 4..
gb - est 3: *
5.. 5..
gb - est 4: *
6.. 6..
gb - est 5: *

```
7: gb_est6:*
8: gb_gsb1:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

No.	Score	Match Length	DB	ID	Description
1	1351	97.9	812	9	CL661282
2	1327	96.2	844	9	CL688009
3	1236	89.6	716	9	CL663485
4	973.5	57.5	830	1	AL669413
5	786	57.0	706		AL669413 AL669413
6	756	54.8	839	9	CL663494
7	614	44.5	1047	9	AY420765
8	607	44.0	1335		PR10152b
9	605	43.8	1047	9	AY420767
					AY420767 Homo sapi
					AY420767 Tetradon
					DR94222 Mus muscu
					AY420767 Mus muscu

10	605	43.8	2550	3	AK077694	Mus muscu
11	604	43.8	1364	3	CR694735	Tetradon
12	603	43.7	870	7	CO163928	FDL-4 ^E
13	601	43.6	739	9	CL685899	PR10425
14	599	43.4	631	8	BH762604	BMBAC330A
15	599	43.4	995	9	AY420766	AY420766
16	594	43.0	1308	3	CH688406	Pan trogl
17	585	42.4	926	5	CR688406	Tetradon
18	578	41.9	890	7	BQ706230	AGENCOURT
19	573	41.5	859	5	CK45875	922961 MA
20	569.5	41.3	967	3	AK006484	BQ232415
21	568	41.2	940	5	BQ937757	AGENCOURT
22	561.5	40.7	997	7	CK264740	EST710818
23	560.5	40.6	780	8	B2573414	B2573414
24	558.5	40.5	867	7	CF205959	msh-2_3093
25	558.5	40.5	885	7	CPF205959	RBR959151
26	556.5	40.3	1235	3	CNS0AB0	CABido005
27	549.5	39.8	876	5	BUL136049	BX816274
28	548.5	39.7	1278	3	CNS02B54	Arabidops
29	547	39.6	742	4	BV434305	BUL36049
30	547	39.6	760	4	BV434312	603124830
31	545	39.5	888	7	CN159313	BV434305
32	542.5	39.3	637	1	AJ722958	BV434312
33	542.5	39.3	818	7	CQ978686	BV432112
34	541	39.2	628	6	CD717720	BV432112
35	541	39.2	761	4	BV434187	BV434187
36	540.5	39.2	1347	8	BZ554193	BV434187
37	537.5	38.9	7	CK269489	BZ554793	
38	533	38.6	778	7	-CQ269489-	pac1-60
39	532.5	38.6	898	7	CN35924	EST15567
40	530.5	38.4	989	7	CW243178	CN359624
41	528	38.3	715	4	CO025828	WS051560
42	527.5	38.2	684	7	CO025828	CW243178
43	527.5	38.2	797	2	BB9468358	WS051560
44	526	38.1	722	6	BB9468358	CW243178
45	526	38.1	CA239168	1	BB9468358	WS051560
731	1	AV881422	ALIGNMENTS	1	AV881422	AV881422

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - nucleic search, using frame_plus_p2n model
 Run on: November 11, 2005, 16:26:32 ; Search time 3123 Seconds
 (without alignments) 4096.118 Million cell updates/sec

Title: US-10-813-549-1
 Perfect score: 1380
 Sequence: 1 MAISIKTPEDIKEKVGAVRL..... CELITLRKDITPAIISHDE 264

Scoring table: BLOSUM62
 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
 Total number of hits satisfying chosen parameters: 9416466
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries:

Command line parameters:
 -MOBLAframe+p2n.model -DEV=xln
 -Q=cgn2_1-USPRO spool/US10813549/runat_03112005_075935_13164/app_query.fasta_1.455
 -DB=GenEmbl -QFORMAT=Fasta -SUFFIX=_rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=20 -MODE=LOCAL
 -OUTFMT=pfo -NORM=ext -THRESHOLD=500 -MINLEN=0 -MAXLEN=000000000
 -USR=USR10813549 @GEN 1 1 @runat 03112005_075935_13164 -NCPU=6 -ICPU=3
 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=10 -LONGLOG
 -DET TIMEOUT=20 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPPOP=6
 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

```
1: GenBml:*
2: gb_ba:*
3: gb_ttg:*
4: gb_in:*
5: gb_on:*
6: gb_dat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1	
LOCUS	EComAp
DEFINITION	E.coli map gene, encoding methionine amino peptidase, complete cds.
ACCESSION	M15106
VERSION	M15106.1
KEYWORDS	map gene; metallo-oligopeptidase; methionine aminopeptidase.
SOURCE	Bacillus coli
ORGANISM	Bacteroides; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE	1 (base 1 to 1197)
AUTHORS	Ben-Bassat,A., Bauer,K., Chang,S.Y., Myambo,K., Boosman,A. and Chang,S.
TITLE	processing of the initiation methionine from proteins: properties of the Escherichia coli methionine aminopeptidase and its gene structure.
JOURNAL	J. Bacteriol. 169 (2), 751-757 (1987)
MEDLINE	87109068
PUBMED	3027045
COMMENT	Original source text: E.coli (strain CM99) DNA, clone PSY174. Draft entry and computer-readable sequence for [1] kindly provided by S.Chang, 01-JUN-1987.
	There are at least two tandem 'map' gene promoters separated by the Small Site at position 150. Two regions of dyad symmetry are located at positions 187-201 and 1018-1037. location/Qualifiers
FEATURES	1. .1197
Source	I02543 Sequence 1

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1380	100.0	1197	1 EComAp
2	1380	100.0	1197	6 E01194
3	1380	100.0	1197	6 I02088 Sequence 1
4	1380	100.0	1197	6 I02543 Sequence 1

OM protein - nucleic search, using frame_plus_p2n model

Run on:

November 11, 2005, 18:29:14 ; Search time 621 seconds

(without alignments) 3515.67 Million cell updates/sec

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Title: US-10-813-549-1
Perfect score: 1380
Sequence: 1 MAISIKTIPDEIKKVRVAGRL.....CBILTRKDDTIPAISSHDE 264

Scoring table: BLOSUM62
Xgapop 10.0 , **Xgapext** 0.5
Ygapop 10.0 , **Ygapext** 0.5
Fgapop 6.0 , **Fgapext** 7.0
Delop 6.0 , **Delext** 7.0

Searched:

9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE:frame+P2N,model -DEV=xlh

-O=/cgn2_1/USPTO_spool/US1013549/runat_03112005_075936_13214/app_query.fasta_1.455

-Db=Published_Applications_NA -QWTF=fastap -SUFFIX=impb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62

-TRANS=human40_cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 ALIGN=20 -NODE=LOCAL -OUTFILE=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USERB=US1013549 @CQN 1.1 480 @runat 03112005_075936_13214

-NCPU=6 -ICLU=3 -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPO=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA,*

1: /cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PCN_NEW_PUB_seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB_seq: *
4: /cgn2_6/ptodata/1/pubpna/us06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/us07_NEW_PUB_seq: *
6: /cgn2_6/ptodata/1/pubpna/PCNTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/us08_PUB_seq: *
8: /cgn2_6/ptodata/1/pubpna/us08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/us09_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/us09_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/us09_PUBCOMB.seq: *
12: /cgn2_6/ptodata/1/pubpna/us09_NEW_PUB_seq: *
13: /cgn2_6/ptodata/1/pubpna/us10_PUB_seq: *
14: /cgn2_6/ptodata/1/pubpna/us10_PUBCOMB.seq: *
15: /cgn2_6/ptodata/1/pubpna/us10B_PUBCOMB.seq: *
16: /cgn2_6/ptodata/1/pubpna/us10C_PUBCOMB.seq: *
17: /cgn2_6/ptodata/1/pubpna/us10D_PUBCOMB.seq: *
18: /cgn2_6/ptodata/1/pubpna/us10E_PUBCOMB.seq: *
19: /cgn2_6/ptodata/1/pubpna/us10F_PUBCOMB.seq: *
20: /cgn2_6/ptodata/1/pubpna/us10G_PUBCOMB.seq: *
21: /cgn2_6/ptodata/1/pubpna/us10H_PUBCOMB.seq: *
22: /cgn2_6/ptodata/1/pubpna/us10I_PUBCOMB.seq: *
23: /cgn2_6/ptodata/1/pubpna/us10J_PUBCOMB.seq: *
24: /cgn2_6/ptodata/1/pubpna/us10K_PUBCOMB.seq: *
25: /cgn2_6/ptodata/1/pubpna/us11A_PUBCOMB.seq: *
26: /cgn2_6/ptodata/1/pubpna/us11B_PUBCOMB.seq: *
27: /cgn2_6/ptodata/1/pubpna/us11C_PUBCOMB.seq: *
28: /cgn2_6/ptodata/1/pubpna/us60_PUBCOMB.seq: *

RESULT 1
US-10-893-671-30/c
; Sequence 30, Application US10893671
; Publication No. US2005064527A1
; GENERAL INFORMATION:
; APPLICANT: LEVY, STUART, et. al.
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PRZ-043
; CURRENT APPLICATION NUMBER: US10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10

ALIGNMENTS

Result No.	Score	Query Match Length	DB ID	Description
SUMMARIES				
C 1	1380	100.0	11204	US-10-893-671-30 Sequence 30, Appli
C 2	937	67.9	640681	Sequence 1, Appli
C 3	911.5	66.1	908766	Sequence 1685, App
C 4	910.5	66.0	807	Sequence 120, Appli
C 5	910.5	66.0	1830121	Sequence 1, Appli
C 6	910.5	66.0	1830121	Sequence 1, Appli
C 7	910.5	66.0	1830121	Sequence 1, Appli
C 8	909.5	65.9	786	Sequence 58, Appli
C 9	909.5	65.9	786	Sequence 60, Appli
C 10	904.5	65.5	807	Sequence 122, Appli
C 11	852	61.7	2731748	Sequence 1, Appli
C 12	799.5	57.9	96109	Sequence 35, Appli
C 13	758.5	55.0	92934	Sequence 21, Appli
C 14	758.5	55.0	2242716	Sequence 1068, App
C 15	619	44.9	855	Sequence 598, Appli
C 16	619	44.9	9025608	Sequence 1, Appli
C 17	617.5	44.7	1735	Sequence 14629,
C 18	614	44.5	2671	Sequence 1, Appli
C 19	614	44.5	2671	Sequence 1, Appli
C 20	614	44.5	2984	Sequence 2189, Ap
C 21	597.5	43.3	3309400	Sequence 2189, Ap
C 22	595	43.1	1125	Sequence 2189, Ap
C 23	594	43.0	1197	Sequence 2189, Ap
C 24	594	43.0	1197	Sequence 2189, Ap
C 25	590	42.8	747	Sequence 2187, Ap
C 26	590	42.8	747	Sequence 1995, Ap
C 27	590	42.8	1747	Sequence 33, Appli
C 28	587	42.5	1605	Sequence 104571,
C 29	574.5	41.6	1116	Sequence 6897, Ap
C 30	571.5	41.4	2287	Sequence 31924, A
C 31	571.5	41.4	3252	Sequence 31927, A
C 32	565.5	41.0	1360	Sequence 452, Appli
C 33	562.5	40.8	1295	Sequence 26939, A
C 34	562.5	40.8	1306	Sequence 48139, A
C 35	552.5	40.0	3088	Sequence 5641, Ap
C 36	550.5	39.9	1534	Sequence 30704, A
C 37	550.5	39.7	1389	Sequence 168772, Sequence 68096, A
C 38	547.5	39.7	1489	Sequence 1, Appli
C 39	539	39.1	120025	Sequence 101599, Sequence 494, App
C 40	531.5	38.9	1525	Sequence 1, Appli
C 41	519	37.6	762	Sequence 1, Appli
C 42	519	37.6	1856	Sequence 1, Appli
C 43	518.5	37.6	943	Sequence 1, Appli
C 44	518.5	37.6	1734	Sequence 1, Appli
C 45	516	37.4	807	Sequence 1, Appli

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 14:27:17 ; Search time 406 Seconds

(without alignments) 3849.290 Million Cell updates/sec

Title: US-10-813-549-1

Perfect score: 1380

Sequence: 1 MAISIKTIPEDIEKMRVAGRL.....CEILTIRKDITPAISHDE 264

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 DelOp 6.0 , Delext 7.0

Searched: 4330206 seqs, 2959370667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODBL=frame+ p2n.model -DEV=x1h
 -O/cgn2_1/USPTO_spol/US10813549/runat_03112005_075935_13156/app_query.fasta_1.455

-DB=Geneseq_16Dec04 -QPMTH=fastap -SUPFix=xmg -MINMATCH=0.1 -LOOPCf=0

-LOOPEXT=0 -UNITSbits -STAR=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdd

-LIST=45 -DOCLIGN=20 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=20

-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEASize=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10813549 @CGN_1 -I 470 @runat_03112005_075935_13156_NCP05 -ICP03

-NO_MMAR -LARGEQUERY -NEQ_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THRBLDS=1 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -FGAPOP=6

Database :

1: N_GeneSeq_16Dec04-*
 2: geneseqn1980s-*
 3: geneseqn2000s-*
 4: geneseqn2001ass-*
 5: geneseqn2001bs-*
 6: geneseqn2002ass-*
 7: geneseqn2002bs-*
 8: geneseqn2003ass-*
 9: geneseqn2003bs-*
 10: geneseqn2003cs-*
 11: geneseqn2003ds-*
 12: geneseqn2004ass-*
 13: geneseqn2004bs-*

RESULT 1

AAN/0151
 ID AAN70151 standard; DNA; 1197 BP.

XX AC AAN70151;

XX DT 09-APR-1991 (first entry)

XX DE Insert of psVCl174 encoding Met-aminopeptidase.

XX KW Met-aminopeptidase; antibody.

XX OS Escherichia coli.

XX FH Key Location/Qualifiers
 FT CDS 219..1013
 /tag= a

XX EP219237-A.

XX PD 22-APR-1987.

XX PF 19-SEP-1986; 86EP-00307242.

XX PR 20-SEP-1985; 85US-00778144.
 PR 06-MAY-1986; 86US-00860330.

XX PA (CETU) CETUS CORP.

ALIGNMENTS

C	6	1137.5	82.4	110000	10	ACF65386_1	Continuation (2 of
C	7	1086	78.7	807	10	ADF02139	Adf02139 Bacterial
C	8	937	67.9	110000	6	ABD92787_2	Continuation (3 of
C	9	911.5	66.1	308766	13	ADT05738	Adt05738 Haemophil
C	10	910.5	66.0	110000	2	ADT42033_17	Adt42033 Haemophil
C	11	910.5	65.9	786	10	ADG73150	Adg73350 Pseudomon
C	12	909.5	65.9	786	10	ADG73152	Adg73352 Pseudomon
C	13	909.5	65.7	861	9	ADA29021	Ada29021 DNA encod
C	14	909.5	65.9	858	11	ABD0567	Abd0567 Pseudomon
C	15	909.5	65.9	963	11	ABD0531	Abd0531 Pseudomon
C	16	909.5	65.9	1176	11	ABD0510	Abd05710 Pseudomon
C	17	904.5	65.5	807	10	ADG73114	Adg73414 Haemophil
C	18	894.5	59.7	861	9	ADA29021	Ada29021 DNA encod
C	19	799.5	57.9	807	12	ADL0291	Adl02971 DNA encod
C	20	799.5	57.9	96109	4	AAF28548	Aaf28548 Genomic f
C	21	798	57.8	975	9	ADA32163	Ada32163 DNA encod
C	22	758.5	55.0	92934	3	AAB81473	Aab81473 N. mening
C	23	758.5	55.0	110000	3	AAB81489_7	Contamination (8 of
C	24	758.5	55.0	172325	3	AAF21613	Aaf21613 Neisseria
C	25	739	53.6	1125	11	ABD15964	Abd15964 Pseudomon
C	26	739	53.6	1248	11	ABD15964	Abd15964 Pseudomon
C	27	739	53.6	1248	11	ABD16453	Abd16453 Pseudomon
C	28	701	50.8	798	10	ADF03390	Adf03390 Bacterial
C	29	682	49.4	2706	2	AVV81790	Avv81790 Granulocy
C	30	618.5	44.8	645	10	ABZ27943	Abz27943 N. gonorr
C	31	618.5	44.7	110000	4	AAB19682_31	Contamination (132 o
C	32	617	44.7	110000	4	AAB19683_31	Contamination (132 o
C	33	616.5	44.7	1625	3	AAC44483	Aac44483 Zea mays
C	34	614.5	44.5	1578	3	AAC43733	Aac43733 Zea mays
C	35	614	44.5	2671	10	ADG4828	Adg4828 Human met
C	36	614	44.5	2671	10	ADG2842	Adg4824 Human sof
C	37	614	44.5	2671	12	ADG18284	Abk80146 cDNA enco
C	38	614	44.5	2984	6	ABK8646	Aaf71155 Corynebac
C	39	597.5	43.3	996	4	AAF71155	Aaf71155 Corynebac
C	40	597.5	43.3	349980	5	AAH68531	Aah68531 C. glutami
C	41	597.5	43.3	349980	5	AAH68530	Aah68530 C. glutami
C	42	595.5	43.2	789	5	AAH67163	Aah67163 C. glutami
C	43	595	43.1	1125	4	ABL23125	Abi23125 Drosophil
C	44	594	43.0	1197	6	ABZ4622	Abz4622 Arabidops
C	45	594	43.0	1419	13	ADR42904	Adr42904 Arabidops

Continuation (2 of

Adf02139 Bacterial

Continuation (3 of

Adt05738 Haemophil

Continuation (18 o

Adg73414 Haemophil

Adg73350 Pseudomon

Abd0567 Pseudomon

Abd0531 Pseudomon

Abd05710 Pseudomon

Adg73414 Haemophil

Adg73350 Pseudomon

Ada29021 DNA encod

Adl02971 DNA encod

Abd15964 Pseudomon

Abd15963 Pseudomon

Adf03390 Bacterial

Adg4828 Human met

Adg4824 Human sof

Abk80146 cDNA enco

Aaf71155 Corynebac

Aah68531 C. glutami

Aah68530 C. glutami

Aah67163 C. glutami

Abi23125 Drosophil

Abz4622 Arabidops

Adr42904 Arabidops

Continuation (28 o

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1380	100.0	1197	1 ANA70151 Insert of Aas70151
2	1380	100.0	1204	4 AAS64245 DNA encoded
3	1330	95.4	846	11 ACH99423 Ach99423 Klebsiell
4	1137.5	82.4	798	10 ACF69733 Photorhab
5	1137.5	82.4	110000	10 ACF67367 Continuation (28 o